Supplementary material

Data description

Table 1: Complete set of attributes in the raw original dataset.

Attribute	Description
indebtedness	Level of indebtedness
credit_amount	Amount of credit
property_value	Property value
loan_to_value	Loan to value
duration	Duration of the loan
studies	Level of studies
poverty_index	Marginalization/poverty index
age	Age
gender	Gender
est_soc_income	Estimated socio-demographic income
value_m2	Value per square meter
est_income	Estimated income
installment	Monthly installment
n_family_unit	Members of the family unit
est_mila_income	Estimated income based on MILA model
p_default	Percentage of defaulted contracts in the last 4 months from those signed during the previous 12 to 24 months
zip_code	ZIP code
municipality	Municipality
economy_level	Level of economy

Table 2: New attributes created during the preprocessing step.

Attribute	Description
zip_code & *nmunicipality*	Bivariate attributed resulting from the concatenation of features *zip_code* and *municipality*
est_soc_income / *est_mila_income*	Univariate attribute resulting from the ratio between features *est_soc_income* and *est_mila_income*
property_value / *installment*	Univariate attribute resulting from the ratio between features *property_value* and *installment*
indebtedness / *loan_to_value*	Univariate attribute resulting from the ratio between features *indebtedness* and *loan_to_value*